

基于动态渐进融合的无人机海上救援目标检测算法

黄绿娥^{1,2}, 于晓伟¹, 鄢化彪¹, 毛玉婷³

(1. 江西理工大学电气工程与自动化学院, 赣州 341000; 2. 多维智能感知与控制江西省重点实验室, 赣州 341000; 3. 江西省生态文明研究院, 南昌 330046)

摘要: 无人机(Unmanned aerial vehicle, UAV)目标检测在海上救援任务中发挥着重要作用。然而, 由于无人机空中拍摄的视角和高度多变, 检测目标存在多尺度变化。此外, 阳光照射海面产生的耀斑会造成误检现象。基于上述问题, 为满足无人机实时目标检测的算法轻量化需求, 本文以YOLOv8n为基准网络, 提出一种基于动态渐进融合的轻量级无人机海上救援目标检测算法(Dynamic progressive fusion YOLO, DPF-YOLO)。首先, 提出轻量级冗余信息提取模块(Redundant information extraction module, RIEM), 通过减少特征图中的冗余信息, 突出关键特征, 避免耀斑误检问题。其次, 提出动态多尺度特征提取模块(Dynamic multi-scale feature extraction module, DMFEM), 通过动态调整感受野大小以适应不同尺度的目标, 增强多尺度特征表达能力。最后, 结合DMFEM模块提出动态渐进融合网络(Dynamic progressive fusion network, DPFNet), 通过渐进式融合结构, 减少非相邻层间不同尺度目标的语义差异, 增强多尺度特征融合效果。DPF-YOLO设计为P2、P3和P4检测层结构以适应海上救援任务中不同尺度的目标, 丰富多尺度信息, 增强对小目标的特征提取。在SeaDronesSee v2数据集上的实验结果表明, DPF-YOLO以仅1.19M的参数量实现了 $mAP_{0.5}=72.2\%$ 的检测精度, 较基准网络YOLOv8n参数量降低60.5%, 召回率提升12.4%, 精度提升8.2%。在VisDrone数据集上的泛化性实验结果表明, DPF-YOLO具有较好的泛化能力。

关键词: 目标检测; 无人机海上救援; 多尺度; 渐进融合; 轻量化模型

中图分类号: TP391.4

文献标志码: A

Object Detection Algorithm for UAV Maritime Rescue Based on Dynamic Progressive Fusion

HUANG Lve^{1,2}, YU Xiaowei¹, YAN Huabiao¹, MAO Yuting³

(1. School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, China; 2. Jiangxi Province Key Laboratory of Multidimensional Intelligent Perception and Control, Ganzhou 341000, China; 3. Jiangxi Research Academy of Ecological Civilization, Nanchang 330046, China)

Abstract: Unmanned aerial vehicle (UAV) object detection plays a crucial role in maritime rescue missions. However, the varying perspectives and altitudes inherent in UAV aerial photography lead to multi-scale variations in object individuals and vessels. Additionally, the glare resulting from sunlight reflecting off the sea surface can cause false detection issues. To address these challenges and meet the lightweight requirements of real-time object detection algorithms for UAVs, this paper proposes a lightweight UAV

基金项目: 国家自然科学基金(62001202); 江西省自然科学基金(20224BAB202036); 江西省教育厅科学技术重点研究项目(GJJ2200805)。

收稿日期: 2025-01-22; **修订日期:** 2025-03-11

maritime rescue object detection algorithm based on dynamic progressive fusion (DPF-YOLO), using YOLOv8n as the baseline network. Firstly, we introduce a lightweight redundant information extraction module (RIEM) that reduces redundant information in feature maps, highlighting key features to mitigate false detections caused by glare. Secondly, we propose a dynamic multi-scale feature extraction module (DMFEM) that dynamically adjusts the receptive field to accommodate objects of varying scales, enhancing multi-scale feature representation capabilities. Finally, by integrating the DMFEM module, we develop a dynamic progressive fusion network (DPFNet). This network employs a progressive fusion structure to reduce semantic differences between non-adjacent layers with objects of different scales, thereby improving multi-scale feature fusion. DPF-YOLO is designed with P2, P3 and P4 detection layer structure to accommodate the object sizes in maritime rescue scenarios, enrich multi-scale information, and enhance feature extraction for small objects. Experimental results on the SeaDronesSee v2 dataset demonstrate that DPF-YOLO achieves a detection accuracy of $mAP_{0.5} = 72.2\%$ with only 1.19 M of parameters. Compared to the baseline network YOLOv8n, DPF-YOLO reduces the number of parameters by 60.5%, increases the recall rate by 12.4%, and improves precision by 8.2%. The generalization experimental results on the VisDrone dataset demonstrate that DPF-YOLO possesses excellent generalization capabilities.

Key words: object detection; unmanned aerial vehicle (UAV) maritime rescue; multi-scale; progressive fusion; lightweight model

引 言

《交通运输部等二十三个部门和单位关于进一步加强海上搜救应急能力建设的意见》中提到:要加强海上人命搜救能力建设,加快构建“陆海空天”一体化感知系统,提升海上目标自动感知和动态跟踪监测能力。海上救援任务的关键在于迅速搜寻到海中遇险人员,并精确估算其相对于救援船的方位。目前,海上救援任务仍需要大量人工参与搜寻,救援效率低^[1]。无人机由于其高机动性、部署灵活、拍摄视野广的优点,在海上救援任务中展现出巨大潜力。将无人机应用于海上救援任务的重点是实现无人机航拍图像的实时目标检测^[2],近年来,能够自适应学习目标特征的深度学习方法发展迅速,已成为目标检测领域的主流技术^[3]。

基于深度学习的目标检测方法主要有双阶段检测方法和单阶段检测方法两类。双阶段检测方法早期代表算法是R-CNN(Region-based convolutional neural network)^[4]和VGG(Visual geometry group)^[5],通过生成候选框并利用卷积神经网络(Convolutional neural network, CNN)进行目标分类。Faster R-CNN^[6]通过引入区域建议网络(Region proposal network, RPN)将候选框和检测合并,进一步提升检测速度和准确性。Mask R-CNN^[7]则在Faster R-CNN基础上新增实例分割功能,实现精准像素级目标掩码。但双阶段检测方法效率低,需要消耗大量的计算资源,不适用于如今复杂的目标检测任务。单阶段检测方法能够在卷积神经网络中同时执行检测和分类两个过程,在实时性方面更有优势。YOLO系列^[8-10]是目前应用最广、实用性最高的单阶段检测方法,能够直接从完整图像中预测边界框位置和类别,提升检测速度,保持较高精度。但在海上救援任务中,通过无人机航拍取得的图像具有独特特点,为传统的YOLO算法带来重大挑战。例如, Jing等^[11]提出的无人机航拍向日葵生长状态检测算法Sunflower-YOLO, Wang等^[12]提出的无人机航拍检测算法YOLOSAR-Lite,均只适用于目标物体大小、形状相似和背景变化均匀的场景,但在无人机海上救援任务中,由于无人机飞行高度的随机性,目标存在尺度、角度和方向变化,上述方法对于特定的应用场景有着显著优势,但在无人机航拍变化多样的场景中具有一定挑战性。

科研人员主要对特征提取阶段进行优化以推动无人机海上救援目标检测算法发展。例如,路径聚合网络(Path aggregation network, PAN)^[13]将浅层语义信息从下向上传输,增强深层语义信息表达。Han等^[14]通过在PAN中添加一条新的分支,在海面目标检测任务中引入多尺度信息,增强多尺度表征

能力。李生辉等^[15]通过在YOLOv5中引入双向特征金字塔网络(Bi-directional feature pyramid network, BiFPN),提升网络对多尺度特征的融合能力。Yang等^[16]将C2f模块和压缩与激励空间注意力(Squeeze-and-excitation spatial attention, SESA)结合设计C2fSESA模块,对不同深度层特征进行聚合,提高对海上小目标的检测精度。Zhang等^[17]基于SeaDroneSee海上救援数据集提出ABT-YOLOv7,通过结合BiFormer模块和上下文解耦模块,增强模型检测小目标的能力。然而,海面经过太阳照射会产生大面积耀斑,海面耀斑在海上救援目标检测任务中会造成误检现象^[18],上述方法均没有解决海面耀斑对目标检测的影响。本文通过改进特征提取模块解决海面耀斑造成的误检问题。

由于无人机固有的飞行特性,海上救援目标检测任务会因无人机的高度、角度变化导致目标尺寸大小多变,影响海上目标视觉特征和整体图像质量。研究人员通过优化网络结构以解决上述问题。例如,Jiang等^[19]通过在特征金字塔网络(Feature pyramid network, FPN)中增加更多的横向连接,以重用每个卷积阶段的特征,实现对不同尺度特征的充分利用。邓天民等^[20]通过设计高效主干网络实现对冗余特征的高效复用,提升模型对不同尺度特征信息的表征能力。邬开俊等^[21]通过设计多分支语义增强模块,实现在多个分支中提取丰富的多尺度特征并构建多尺度特征之间的联系,从而在特征融合传递信息时防止重要特征信息丢失。Jiang等^[22]结合双向密集特征金字塔网络(Bidirectional dense FPN, BDFPN),通过扩展特征金字塔网络的尺度并引入跳跃连接,减少冗余信息的干扰,实现高效的多尺度特征融合。然而,上述方法忽略了随着网络深度加深造成的特征退化现象,本文通过设计新型颈部网络以减少不同尺度目标语义差异,减轻特征退化现象。

针对上述问题,综合考虑海面耀斑、检测精度和模型参数量等指标,提出基于动态渐进融合的轻量级无人机海上救援目标检测算法(Dynamic progressive fusion YOLO, DPF-YOLO)。主要贡献如下:首先,针对海面耀斑造成的误检问题,提出轻量级冗余信息提取模块(Redundant information extraction module, RIEM),轻量级RIEM模块通过减少特征图中空间和通道纬度的冗余信息,减少冗余信息计算,避免耀斑干扰并降低参数量。其次,针对无人机海上救援目标多尺度问题,提出动态多尺度特征提取模块(Dynamic multi-scale feature extraction module, DMFEM),通过动态调整感受野大小适应不同尺度的目标,突出不同尺度的特征信息。最后,针对特征退化现象,结合DMFEM模块提出动态渐进融合网络(Dynamic progressive fusion network, DPFNet),通过引入针对非相邻层特征的融合路径,减少非相邻层不同尺度特征的语义差异,保留更多的多尺度特征信息。此外,将检测层设计为P2、P3和P4三层结构,以适应海上救援目标尺寸,丰富多尺度信息。

1 DPF-YOLO算法

DPF-YOLO算法的整体网络结构如图1所示。主干中提出轻量级RIEM模块进行高效特征提取,通过对通道和空间维度的冗余信息进行压缩重构,避免耀斑的干扰并降低参数量。提出DPFNet替换YOLOv8原本的颈部网络,通过渐进式融合结构促进非相邻层级不同尺度特征的交互融合,丰富多尺度特征信息,减轻目标特征衰减。DPFNet中结合本文提出的DMFEM模块增强多尺度特征提取,通过动态调整感受野大小适应不同尺度的目标特征,突出关键信息。卷积(Convolution, Conv)在卷积神经网络中用于提取图像特征,输入Conv的图像首先经过二维卷积(2D convolution, Conv2d)进行处理,之后经过二维批量归一化操作(Batch normalization for 2D data, BatchNorm2d)进行归一化处理,最后由激活函数SiLU(Sigmoid linear unit)处理。主干中使用的SPPF(Spatial pyramid pooling-fast)是一种空间金字塔池化方法,通过快速池化减少计算复杂度。动态渐进融合网络中使用自适应空间特征融合(Adaptive spatial feature fusion, ASFF)操作^[23]来动态调整不同尺度特征图的贡献度,其中ASFF-2和ASFF-3分别表示在两个不同的融合阶段中使用ASFF操作。头部设计使用解耦头(Decoupled head)对特征图进行最后处理。解耦头中,边界框损失函数(Bounding box loss, Bbox.Loss)用于计算预测边界框与真实边界框之间差异,分类损失函数(Classification loss, Cls.Loss)用于计算预测类别与真实类别之间的差异。

为适应无人机视角下的海上救援目标尺寸,DPF-YOLO算法采用P2、P3和P4检测层结构,3个检测层输出特征图的尺寸分别为160像素×160像素、80像素×80像素和40像素×40像素。P2检测层具有更高的分辨率,本文实验表明,P2检测层的引入能够提升网络对小目标的检测性能。此外,将YOLOv8中的P5检测层去除,减少针对大目标的结构特征,降低整体复杂度。

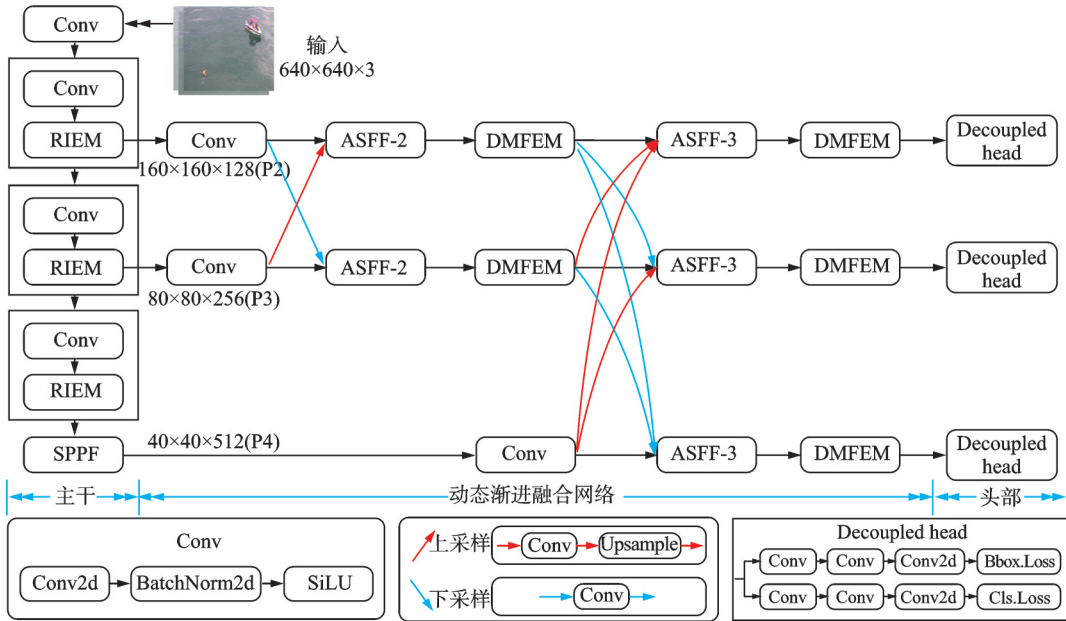


图1 DPF-YOLO整体网络结构

Fig.1 Overall network structure of DPF-YOLO

1.1 冗余信息提取模块

C2f模块瓶颈(Bottleneck)中通过连续使用卷积模块以捕获丰富的特征信息,但该方式会引入部分冗余信息和大量卷积计算,使特征提取过程易受背景信息(如阳光照射海面产生的耀斑)的干扰,并增加模块参数量。为解决以上问题,本文提出RIEM模块。RIEM模块详细结构及其与C2f模块的热力图可视化对比如图2所示。RIEM模块能够减少空间和通道维度的冗余信息,降低对复杂背景等无用信

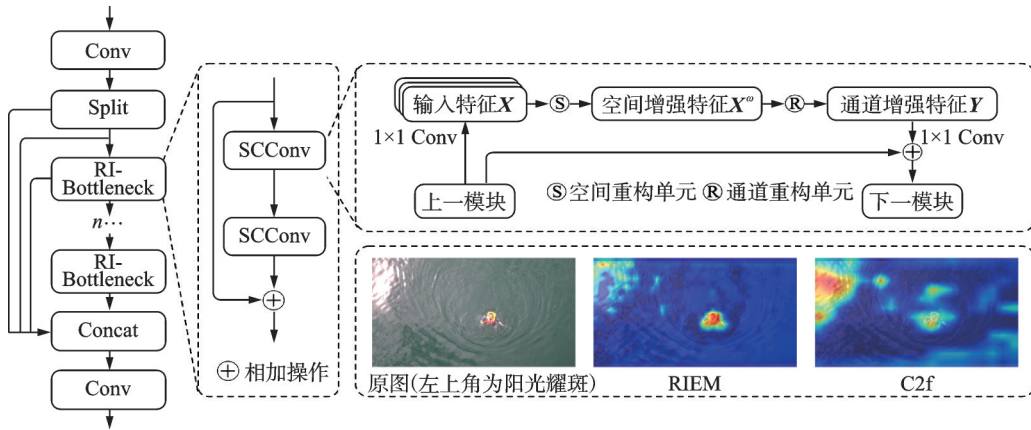


图2 RIEM模块详细结构和热力图可视化对比

Fig.2 Detailed structure of the RIEM module and comparison of heatmap visualization

息的关注度,并通过减少冗余信息实现对特征图的压缩,降低参数量。由图2中的热力图对比可知,经C2f操作后的特征响应分散在非目标区域的阳光耀斑和海面,相比之下,经RIEM操作后的特征响应集中在目标物体,海面耀斑区域被有效抑制,说明RIEM模块能够增强网络的特征提取能力,避免无用的背景信息干扰。

RIEM模块中的RI-Bottleneck使用空间和通道重构卷积(Spatial and channel reconstruction convolution, SCConv)^[24]进行设计,输入SCConv的特征 X 首先经过空间重构单元(Spatial reconstruction unit, SRU)获得空间增强特征 X^ω ,之后经过通道重构单元(Channel reconstruction unit, CRU)获得通道增强特征 Y ,实现方式如图3所示。

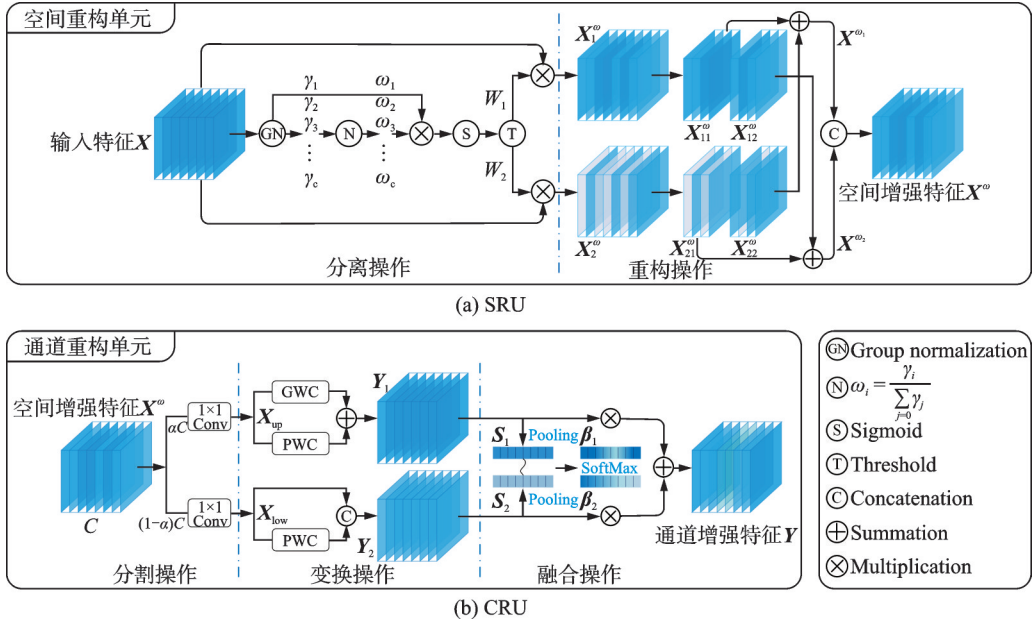


图3 SRU和CRU的实现方式

Fig.3 Implementation methods of SRU and CRU

SRU将输入特征 X 进行先分离后重构操作。首先,分离操作将空间信息中信息丰富的特征图与信息较少的特征图分别分组,采用组归一化(Group normalization, GN)层中的可训练参数 γ 衡量每个批次和通道上的空间像素方差, γ 的大小与空间像素变化程度相关,其值越大代表特征图含有越丰富的空间信息。海面上的耀斑区域通常呈现均匀高亮度分布,空间梯度变化较小,对应的 γ 值较低,而船只、游泳者等目标具有显著的边缘结构和纹理特征,对应的 γ 值较高,含有更丰富的空间信息。该过程计算公式为

$$W_\gamma = \{w_i\} = \frac{\gamma_i}{\sum_{j=1}^c \gamma_j} \quad i, j = 1, 2, \dots, C \quad (1)$$

式中: $W_\gamma \in \mathbb{R}^C$ 为归一化相关权重。随后,使用Sigmoid函数将组归一化权重 W_γ 重新加权的特征图映射在范围(0,1)之间,阈值(Threshold)设置为0.5。将高于阈值的权重设置为1,记为信息权重 W_1 ;低于阈值的权重设置为0,记为信息权重 W_2 。再将输入特征 X 分别与 W_1 、 W_2 相乘,生成对应的加权特征 X_1^ω 与 X_2^ω 。 X_1^ω 具有丰富的空间信息,而 X_2^ω 几乎不含空间信息,视为冗余特征。采用交叉重构操作结合两种不同的信息特征,以加强 X_1^ω 、 X_2^ω 之间的信息流动。最后,将两个经过重构的特征 X^{ω_1} 、 X^{ω_2} 进行拼接,得到空间增强特征 X^ω 。整个重构过程计算公式为

每个分支的通道数为原来的一半,每个分支上使用 1×1 卷积对分割后的特征图进行预处理。右侧分支中使用感受野注意力卷积(Rceptive-field attention convolutional, RFACnv)层动态捕获感受野空间特征,之后将输出的特征图与左侧分支进行拼接,最后进行一次卷积操作,得到输出特征。

RFACnv使用单独的窗口捕获感受野内的全局信息,通过 3×3 组卷积(Group conv)将提取的全局信息转换为感受野空间特征,每个感受野区域通过平均池化操作实现聚合。之后使用组卷积运算来进行信息交互,最后使用SoftMax函数强调每个特征在其各自感受野中的重要性。该过程计算公式为

$$F = A_{rf} \times F_{rf} = \text{SoftMax}(g^{1 \times 1}(\text{AvgPool}(X))) \times \text{ReLU}(\text{Norm}(g^{k \times k}(X))) \quad (8)$$

式中: $g^{i \times i}$ 为具有 $i \times i$ 空间范围的组卷积运算, k 为卷积核大小, Norm为归一化过程, X 为输入特征图,注意力图 A_{rf} 与转换后的感受野空间特征 F_{rf} 相乘得到 F 。

1.3 动态渐进融合网络

在设计无人机海上救援目标检测网络时,关键在于实现跨多个尺度的特征融合与交互,以确保足够的检测精度。如图5所示, YOLOv8颈部采用自上而下的FPN网络和自下而上PAN网络将浅层与深层语义信息相结合, BiFPN网络^[25]在FPN-PAN网络双向特征金字塔结构的基础上,引入可学习的权重来平衡不同输入特征的重要性。然而,上述网络专注于融合相邻层级特征,在融合非相邻层级特征时会增加不同尺度特征的差异性,加速目标特征的衰减。为此,本文提出DPFNet。DPFNet设计为渐进式融合结构,通过引入对非相邻层特征的融合路径,减少非相邻层不同尺度特征的语义差异,保留更多的多尺度特征信息。

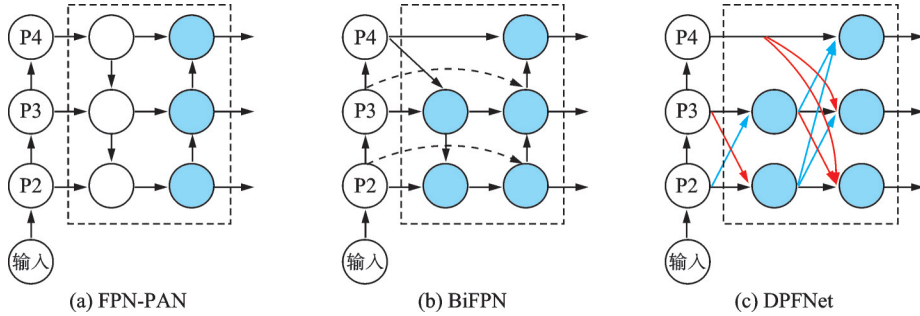


图5 FPN-PAN、BiFPN和DPFNet网络结构对比图

Fig.5 Network structure comparison of FPN-PAN, BiFPN and DPFNet

DPFNet详细结构如图6所示,共分为两个融合阶段:一阶段融合浅层和中间层特征,二阶段继续融合深层特征,实现从浅层特征到深层特征的渐进式融合。每个融合阶段后使用DMFEM模块进行特征提取,通过动态调整感受野大小突出不同尺度的目标特征,减轻来自不同层级冗余信息的影响,避免淹没有效特征。在融合过程中,斜向下的红色虚线箭头代表上采样过程,由卷积与Upsample模块构成;斜向上的蓝色箭头代表下采样过程,由卷积构成。

为减少不同目标相互矛盾的信息对特征融合的影响,引入ASFF操作,通过为不同层次的特征分配不同的空间权重,增强关键信息显著性。ASFF实现如下

$$y_{ij}^l = \alpha_{ij}^l \cdot x_{ij}^{1 \rightarrow l} + \eta_{ij}^l \cdot x_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \cdot x_{ij}^{3 \rightarrow l} \quad (9)$$

式中: α_{ij}^l , η_{ij}^l 和 γ_{ij}^l 表示在第 l 层的3层特征的空间权重, $\alpha_{ij}^l + \eta_{ij}^l + \gamma_{ij}^l = 1$ 。 $x_{ij}^{n \rightarrow l}$ 表示在位置 (i, j) 处,从第 n 层到第 l 层的特征。

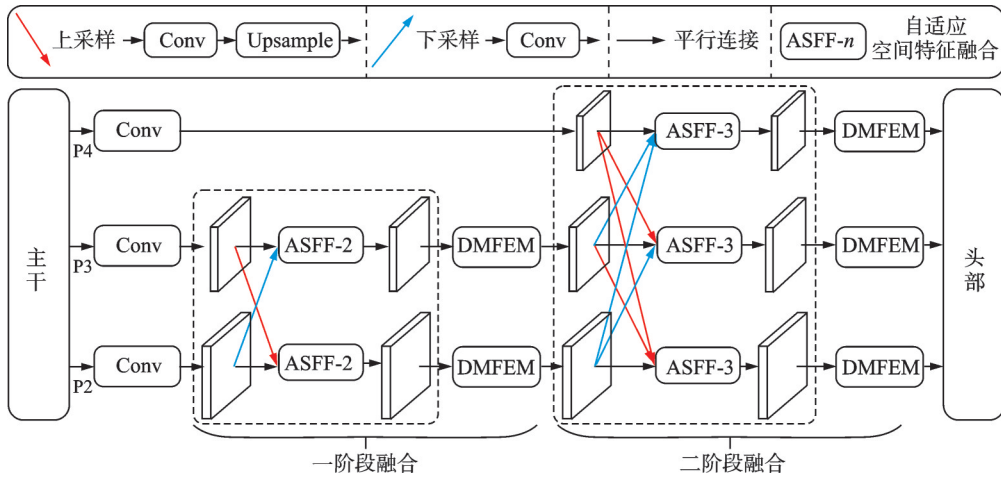


图6 DPFNet网络结构

Fig.6 Network structure of DPFNet

2 实验结果与分析

2.1 数据集

实验采用 SeaDronesSee v2^[26]海上救援数据集,共有 14 227 张由无人机从不同高度和角度拍摄的海上救援图像,其中训练集、验证集和测试集分别包含 8 930、1 547 和 3 750 张图片。该数据集人工标注了 5 个有效目标类别:游泳者、船只、摩托艇、救生设备和浮标,5 个类别的数量分布如图 7(a)所示。数据集各目标的边界框宽高分布如图 7(b)所示,颜色越深表示处于该宽高的目标越多,可以看出目标主要分布在宽高较小的区域,大宽高区域的目标分布较少且零散,体现无人机海上救援图像主要由小尺寸目标组成,且尺度变化较大。

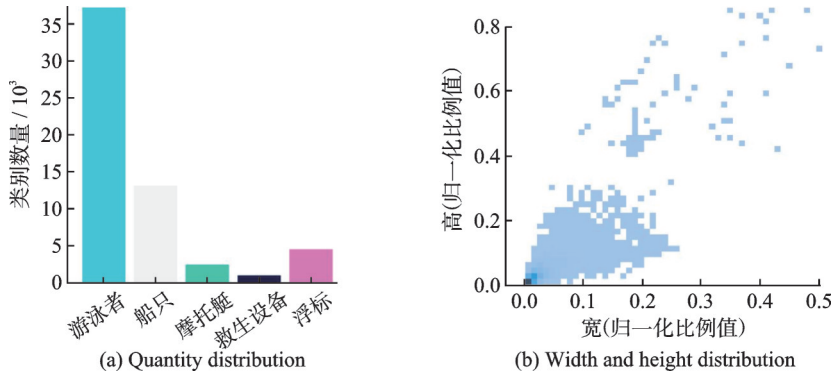


图7 SeaDronesSee v2数据集中人工标注物体的相关信息

Fig.7 Relevant information about manually annotated objects in the SeaDronesSee v2 dataset

2.2 实验配置

模型训练所使用的实验平台和超参数配置如表 1 所示,所有实验在模型训练过程中均未使用预训练权重。为满足部署在无人机边缘设备中的需求,实现高效的目标识别和推理能力,需针对小参数量、低内存占用和高推理速度进行模型优化,因此选择 YOLOv8n 作为基准模型。YOLOv8n 限制了一定的网络宽度和深度,但保留 YOLOv8 系列的核心特性,最适合部署于资源受限的无人机边缘设备中。

表1 实验环境与超参数配置

Table 1 Experimental environment and hyperparameter configuration

实验参数	环境配置	超参数	配置
操作系统	Ubuntu 22.04.4 LTS	输入图像大小	640像素×640像素
CPU	Intel Xeon Platinum 8352V	批量大小	64
GPU	NVIDIA RTX 4090, 24 GB	设定轮数	200
内存	90 GB	早停机制	20
Python	3.10.14	优化器	SGD
Ultralytics YOLO	YOLOv8.1.19	学习率	0.01
深度学习架构	Pytorch2.4.0, CUDA12.4	关闭马赛克数据增强	最后10轮

2.3 评价指标

选用召回率(R)、平均精度(AP)和均值平均精度(mAP,包含IOU阈值为0.5和0.50~0.95时的mAP,分别记为 $mAP_{0.5}$ 和 $mAP_{0.5;0.95}$)作为模型性能和检测精度的评价指标,采用模型参数量(P)、模型规模量(M)、浮点运算总量(Floating point operations per second, FLOPs)和每秒处理速度作为模型复杂度和检测速度的评价指标。计算公式分别为

$$R = \frac{TP}{TP + FP} \times 100\% \quad (10)$$

$$AP = \int_0^1 P(R) dR \times 100\% \quad (11)$$

$$mAP = \frac{1}{N} \sum_{n=1}^N AP_n \times 100\% \quad (12)$$

式中: $P(R)$ 为精度和召回率曲线,TP(True positive)表示正确预测为正类别的数量,FP(False positive)表示错误预测为正类别的数量, N 表示检测类别数量,实验使用的SeaDronesSee v2数据集中 $N=5$ 。

2.4 动态渐进融合网络消融实验

为验证DMFEM模块放置DPFNet不同位置时对多尺度目标检测的贡献程度,在SeaDronesSee v2数据集上进行3组实验。3组实验均设置为P2、P3、P4检测层配置,采用YOLOv8的C2f主干结构,在颈部引入DPFNet进行消融。设计如下:DPFNet-1实验将DMFEM模块插入ASFF-2之后,二阶段融合使用卷积模块;DPFNet-2实验在一阶段融合使用卷积模块,将DMFEM模块插入ASFF-3之后;DPFNet-3实验在两组ASFF操作之后都插入DMFEM模块。实验结果如表2所示,其中加粗数据为该列最佳值。DPFNet-3实验的 $mAP_{0.5}$ 最高,且在“摩托艇、救生设备、浮标”这3类目标中取得最佳AP值。在召回率中,DPFNet-3实验同样取得最高值,说明该实验能够被正确预测的样本数最多。由此可知,引入两组DMFEM模块的DPFNet检测效果最优,虽然参数量较其他两组实验略微增加 0.16×10^6 、 0.04×10^6 ,但在小幅增加参数量的情况下,检测速度仍达到242.5帧/s,满足实时性要求。

表2 DPFNet消融实验结果对比

Table 2 Comparison of ablation experiment results of the DPFNet

模型	$mAP_{0.5}/\%$	AP/ $\%$					$R/\%$	$P/10^6$	速度/ (帧·s ⁻¹)
		游泳者	船只	摩托艇	救生设备	浮标			
DPFNet-1	69.6	76.1	95.8	82.9	27.7	65.6	64.0	1.19	246.1
DPFNet-2	70.7	75.5	95.4	82.3	33.0	67.1	66.0	1.31	243.2
DPFNet-3	72.1	76.0	95.5	87.3	35.1	66.8	70.0	1.35	242.5

2.5 检测层消融实验

为验证不同检测层组合对检测性能的影响,在SeaDronesSee v2数据集上进行3组实验。设计如下:第1组为P3、P4、P5检测层实验(P345),第2组为在第1组实验的基础上加入P2检测层的实验(P2345),第3组为在第2组实验基础上去除P5检测层的实验(P234)。实验结果如表3所示,其中加粗数据为该列最佳值。综合考量5个性能指标,虽然使用P2345检测层的实验 $mAP_{0.5}$ 最高,为72.8%,但参数量高达 5.06×10^6 ,检测速度也有明显下降,以牺牲参数量和检测速度带来的 mAP 精度提升,与轻量化和实时性需求相违背。而使用P2、P3、P4检测层的实验,仅以 1.35×10^6 的参数量达到了72.1%的精度,相较第1组实验 $mAP_{0.5}$ 提升6.0%,参数量仅为第2组实验的26.6%,检测速度也是3组实验的最佳值。由此可知,选用P2、P3、P4检测层能够在提升检测精度的同时满足轻量化和实时性需求。

表3 检测层消融实验结果对比

Table 3 Comparison of detection layer ablation experiment results

模型	$mAP_{0.5}/\%$	AP/ $\%$					$R/\%$	$P/10^6$	速度/ (帧·s ⁻¹)
		游泳者	船只	摩托艇	救生设备	浮标			
P345	66.1	71.2	94.1	84.6	20.8	59.7	60.2	4.38	236.1
P2345	72.8	77.3	96.7	86.5	36.2	67.6	69.0	5.06	228.3
P234	72.1	76.0	95.5	87.3	35.1	66.8	70.0	1.35	242.5

2.6 模块有效性实验

为验证轻量化RIEM模块、DMFEM模块对网络性能提升的有效性,选取YOLOv8中使用的C2f模块、YOLOv11中使用的C3k2模块^[27]在SeaDronesSee v2数据集上进行6组对比实验。设计如下:在DPF-YOLO网络的主干中分别插入C2f、C3k2和RIEM模块进行比较,在颈部DPFNet中分别集成C2f、C3k2和DMFEM模块进行比较。实验结果如表4所示,相较于C2f、C3k2模块,RIEM模块通过减少特征图中的冗余信息,使网络的 $mAP_{0.5}$ 分别提升1.6%、1.1%,召回率分别提升1.9%、3.9%,参数量分别减少 0.16×10^6 、 0.06×10^6 ,检测速度略微下降。传统模块捕获特征的能力容易受固定感受野大小的限制,而DMFEM模块通过动态调整感受野大小适应不同尺度的目标,使网络 $mAP_{0.5}$ 分别提升2.9%、2.3%,召回率分别提升6.2%、6.2%,参数量略微增加 0.03×10^6 、 0.09×10^6 ,满足轻量化与实时性需求。

表4 RIEM模块、DMFEM模块有效性实验结果对比

Table 4 Comparison of experimental results on the effectiveness of RIEM and DMFEM modules

模型	主干	颈部	$mAP_{0.5}/\%$	$mAP_{0.5:0.95}/\%$	$R/\%$	$P/10^6$	速度/(帧·s ⁻¹)
(1)	C2f	DPFNet(C2f)	69.2	38.8	63.8	1.32	249.8
(2)	RIEM	DPFNet(C2f)	70.8	39.7	65.7	1.16	245.4
(3)	C2f	DPFNet(DMFEM)	72.1	40.3	70.0	1.35	242.5
(4)	C3k2	DPFNet(C3k2)	69.3	39.1	63.4	1.16	251.2
(5)	RIEM	DPFNet(C3k2)	70.4	39.4	67.3	1.10	243.6
(6)	C3k2	DPFNet(DMFEM)	71.6	40.1	69.6	1.25	247.2

2.7 改进点消融实验

为验证轻量化RIEM模块、DPFNet、P2检测层对网络性能贡献,在SeaDronesSee v2数据集上从两个方向设计8组消融实验。首先在基准模型YOLOv8n基础上,每个实验仅加入一个改进点,以验证

每个改进点对基准模型的影响;其次将每两个改进点组合后加入基准模型中,以验证每两个改进点组合对DPF-YOLO的影响。实验结果如表5所示,其中“√”表示加入相应改进点进行训练。由表5可知,模型(1)的mAP与召回率较YOLOv8n均有提升,参数量和规模量分别减少 0.39×10^6 和0.74 MB,检测速度下降5.6帧/s,mAP提升得益于RIEM模块能够减少特征提取过程的冗余信息,检测速度下降原因主要是RIEM模块使特征提取过程更关注于关键信息,导致推理速度略微下降。模型(2)的mAP_{0.5}和mAP_{0.5:0.95}较YOLOv8n分别提升2.6%和1.2%,召回率提升1.8%,mAP提升得益于DPFNet的渐进融合结构,减少不同尺度特征的语义差异,集成的DMFEM模块通过动态调整感受野适应不同尺度目标,参数量在DPFNet与P2检测层结合使用时会显著降低。模型(3)的mAP_{0.5}和mAP_{0.5:0.95}分别提升4.9%和1.7%,参数量和规模量分别减少 1.98×10^6 和3.88 MB,mAP提升得益于P2检测层连接的浅层网络能够保留更多的细节信息,增强小目标特征表达,同时引入多尺度信息,增强网络对多尺度目标的适应能力。由模型(4)到模型(6)的实验可知,DPFNet与P2检测层结合使用时对网络性能的贡献度最大,mAP_{0.5}值相较于基准网络提升8.6%,参数量相较于单独使用DPFNet的模型(2)减少 3.04×10^6 ,改善DPFNet造成的参数量增加问题。DPF-YOLO相较于YOLOv8n,mAP_{0.5}和mAP_{0.5:0.95}分别提升8.7%和3.2%,召回率提升12.4%,参数量与规模量分别降低 1.82×10^6 和3.39 MB,仅为YOLOv8n的39.5%和45.7%,浮点运算总量FLOPs为 8.3×10^9 ,检测速度达到246.8帧/s,满足轻量化与实时性需求。

表5 改进点消融实验结果对比

Table 5 Comparison of improved point ablation experiment results

模型	RIEM	DPFNet	P2 检测层	mAP _{0.5} /%	mAP _{0.5:0.95} /%	R/%	P/ 10^6	M/MB	FLOPs/ 10^9	速度/ (帧·s ⁻¹)
YOLOv8n				63.5	37.5	58.4	3.01	6.25	8.1	265.0
(1)	√			64.2(+0.7)	37.9	58.7	2.62	5.51	7.0	259.4
(2)		√		66.1(+2.6)	38.7	60.2	4.38	9.14	8.6	252.3
(3)			√	68.4(+4.9)	39.2	61.8	1.03	2.37	9.5	262.1
(4)		√	√	72.1(+8.6)	40.3	70.0	1.35	3.14	9.7	242.5
(5)	√		√	71.2(+7.7)	40.0	64.9	0.88	2.10	8.2	250.5
(6)	√	√		67.1(+3.6)	38.9	61.2	3.99	8.38	7.9	247.9
DPF-YOLO	√	√	√	72.2(+8.7)	40.7	70.8	1.19	2.86	8.3	246.8

DPF-YOLO与YOLOv8n在SeaDronesSee v2数据集上对5类目标检测的AP如图8所示。可见,相较于YOLOv8n,DPF-YOLO对游泳者、救生设备、浮标3个类别AP提升明显,分别提升10.1%、13.6%、15.1%,对船只、摩托艇AP分别提升4.2%、0.4%。结果表明DPF-YOLO对小目标、多尺度目标检测精度提升明显。

2.8 对比实验

为验证DPF-YOLO在无人机海上救援任务中的检测性能,选取用于多尺度目标和复杂场景检测的PP-PicoDet-L^[28]、SuperYOLO^[29]、YoloOW^[30]和MARF-PNet^[31]算法及其他先进的无人机海上救援目标检测算法在SeaDronesSee v2数据集上进行实验比较。

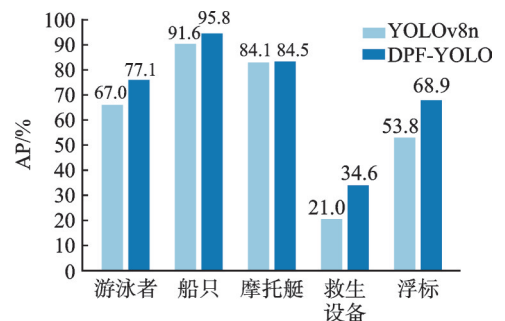


图8 DPF-YOLO与YOLOv8n对5类目标检测的AP对比

Fig.8 Comparison of AP for five-class object detection between DPF-YOLO and YOLOv8n

实验结果如表6所示,加粗数据为该列最优值。由表6可见,在 $mAP_{0.5}$ 方面,DPF-YOLO比YOLOv11n高出8.8%,比两阶段检测算法Faster R-CNN高出14.5%,说明DPF-YOLO在检测精度方面有相当不错的表现。DPF-YOLO不仅在检测精度上超越了对比算法,还在参数量方面占优。DPF-YOLO的参数量仅为 1.19×10^6 ,在 $mAP_{0.5}$ 精度上超越了参数量为 43.67×10^6 的YOLOv8l,综合性能领先于其他先进算法。

综合以上分析,DPF-YOLO在无人机海上救援任务中表现出色,具备轻量化优势的同时仍然能够取得较高的检测精度,且能够避免海面耀斑对检测造成的影响。

YOLOv8n和DPF-YOLO的可视化对比如图9所示。可见,DPF-YOLO的多尺度识别能力更强,漏检率更低。对于由阳光照射海面产生的耀斑,YOLOv8n将耀斑错误识别为“船只”,DPF-YOLO则不存在此错误识别现象。

表6 DPF-YOLO与其他先进算法在SeaDronesSee v2数据集上检测性能对比

Table 6 Comparison of detection performance between DPF-YOLO and other advanced algorithms on the SeaDronesSee v2 dataset

模型	$mAP_{0.5}/\%$	$mAP_{0.5:0.95}/\%$	$P/10^6$	速度/ (帧·s ⁻¹)
Faster R-CNN ^[6]	57.7	34.1	43.10	—
RT-DETR	71.3	40.9	20.18	177.6
YOLOv5n	61.6	36.0	2.51	269.1
YOLOv6n	62.4	37.2	4.25	230.7
YOLOv8n	63.5	37.5	3.01	265.0
YOLOv8s	64.7	38.9	11.10	219.9
YOLOv8m	66.7	41.7	25.90	191.0
YOLOv8l	66.5	41.8	43.67	156.3
YOLOv9t	65.2	38.6	1.97	243.4
YOLOv10n	66.3	39.5	2.70	270.8
YOLOv11n	63.4	37.8	2.58	254.7
PP-PicoDet-L ^[28]	65.6	38.1	3.30	—
SuperYOLO ^[29]	71.2	40.1	7.70	227.5
YoloOW ^[30]	65.4	—	39.30	65.3
MARFPNet ^[31]	48.9	27.0	13.03	—
DPF-YOLO(ours)	72.2	40.7	1.19	246.8

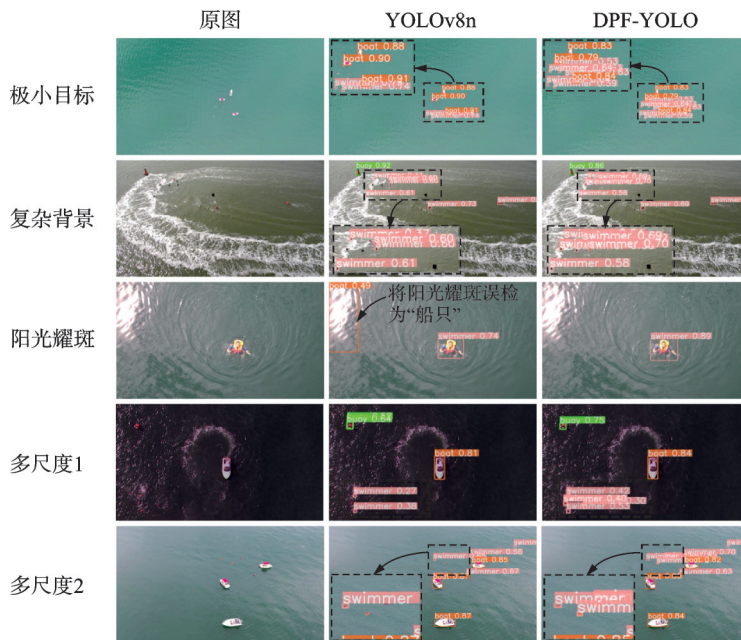


图9 YOLOv8n和DPF-YOLO在SeaDronesSee v2数据集上的可视化对比

Fig.9 Visual comparison of YOLOv8n and DPF-YOLO on the SeaDronesSee v2 dataset

2.9 泛化性实验

为验证 DPF-YOLO 的泛化性和鲁棒性,选用由中国天津大学发布的 VisDrone^[32] 无人机航拍图像目标检测数据集进行泛化性实验。该数据集共有 8 599 张图片,从中国相隔数千公里的 14 个不同城市拍摄,覆盖包括不同的时间、目标、密度、天气等多样场景,目标尺度差异大,适合用于验证算法的泛化能力。按照相同的环境配置进行训练,实验结果如表 7 所示,加粗数据为该列最佳值。DPF-YOLO 的检测精度指标 $mAP_{0.5}$ 相对于基准算法 YOLOv8n 提升 5.3%,参数量仅为之前的 39.5%。综合分析实验结果可知,DPF-YOLO 在陌生数据集上体现出较高的准确度和较强的泛化能力,在面对尺度差异大、小目标众多的数据集时更有优势。

表 7 DPF-YOLO 与其他先进算法在 VisDrone 数据集上检测性能对比

Table 7 Comparison of detection performance between DPF-YOLO and other advanced algorithms on the VisDrone dataset

模型	$mAP_{0.5}/\%$	$mAP_{0.5,0.95}/\%$	$P/10^6$	模型	$mAP_{0.5}/\%$	$mAP_{0.5,0.95}/\%$	$P/10^6$
Faster R-CNN ^[6]	21.8	—	43.10	YOLOv9t	35.5	21.0	1.97
YOLOv5n	27.7	14.4	2.51	YOLOv10n	33.5	19.6	2.70
YOLOv5s	33.4	18.6	7.04	GELAN-T	36.0	21.4	1.88
YOLOv5m	37.7	21.6	20.90	PP-PicoDet-L ^[28]	34.2	20.9	3.30
YOLOv6s	38.1	23.0	16.30	VC-YOLO ^[33]	28.4	—	17.70
YOLOv7-tiny	37.7	19.8	6.03	MPE-YOLO ^[34]	37.0	—	4.40
YOLOv8n	33.9	19.7	3.01	DPF-YOLO(ours)	39.2	23.4	1.19

3 结束语

本文提出一种轻量级无人机海上救援目标检测算法 DPF-YOLO。针对小目标和多尺度目标,提出轻量级 RIEM 模块和 DMFEM 模块,前者通过减少特征图中的冗余信息,在降低模块参数量的同时避免海面耀斑对小目标识别的干扰,后者通过动态调整感受野,增强多尺度特征表达,提高对关键特征的关注能力。针对特征衰减问题,提出集成 DMFEM 模块的 DPFNet,由浅入深逐层提取图像特征,减少非相邻层特征的语义差异,增强细节特征表达能力。检测层结构设计为 P2、P3、P4 三层结构,丰富多尺度信息,增强对小目标的特征提取能力。在 SeaDronesSee v2 数据集上实验表明,DPF-YOLO 以仅 1.19×10^6 的参数量取得 72.2% 的 $mAP_{0.5}$ 精度,相较 YOLOv8n 参数量降低 60.5%, $mAP_{0.5}$ 提升 8.2%。与其他先进的海上救援目标检测算法相比,DPF-YOLO 具有较大优势。目前算法仍存在改进空间,例如对训练样本数量极少的目标精度提升有限,未来研究将通过样本扩充、数据增强等方法进一步提高小样本目标的检测精度,并进一步研究在零样本条件下的目标检测。

参考文献:

- [1] YANG T, JIANG Z, SUN R, et al. Maritime search and rescue based on group mobile computing for unmanned aerial vehicles and unmanned surface vehicles[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(12): 7700-7708.
- [2] ZHANG R, NEWSAM S, SHAO Z, et al. Multi-scale adversarial network for vehicle detection in UAV imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 180: 283-295.
- [3] 周璇, 葛琦, 邵文泽. 高分辨率特征增强的无人机航拍小目标检测[J]. *数据采集与处理*, 2024, 39(4): 908-921.
ZHOU Xuan, GE Qi, SHAO Wenzhe. Small target detection in UAV aerial imagery via high-resolution feature enhancement[J]. *Journal of Data Acquisition and Processing*, 2024, 39(4): 908-921.

- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014: 580-587.
- [5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10). <https://doi.org/10.48550/arXiv.1409.1556>.
- [6] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [7] LI K, WAN G, CHENG G, et al. Object detection in optical remote sensing images: A survey and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 159: 296-307.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [9] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE, 2023: 7464-7475.
- [10] WANG C Y, YE H I, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [EB/OL]. (2024-03-21). <https://doi.org/10.48550/arXiv.2402.13616>.
- [11] JING R, NIU Q, TIAN Y, et al. Sunflower-YOLO: Detection of sunflower capitula in UAV remote sensing images[J]. European Journal of Agronomy, 2024, 160: 127332.
- [12] WANG H, SHI J, KARIMIAN H, et al. YOLOSAR-Lite: A lightweight framework for real-time ship detection in SAR imagery[J]. International Journal of Digital Earth, 2024, 17(1): 2405525.
- [13] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 8759-8768.
- [14] HAN Y, YANG X, PU T, et al. Fine-grained recognition for oriented ship against complex scenes in optical remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-18.
- [15] 李生辉, 李晓飞, 宋璋晗, 等. 基于改进YOLOv5的船舶多尺度SAR图像检测算法[J]. 数据采集与处理, 2024, 39(1): 120-131.
- LI Shenghui, LI Xiaofei, SONG Zhanghan, et al. Multi-scale SAR ship detection algorithm based on improved YOLOv5[J]. Journal of Data Acquisition and Processing, 2024, 39(1): 120-131.
- [16] YANG Z, YIN Y, JING Q, et al. A high-precision detection model of small objects in maritime UAV perspective based on improved YOLOv5[J]. Journal of Marine Science and Engineering, 2023, 11(9): 1680.
- [17] ZHANG Y, YIN Y, SHAO Z. An enhanced target detection algorithm for maritime search and rescue based on aerial images [J]. Remote Sensing, 2023, 15(19): 4818.
- [18] LYU H, SHAO Z, CHENG T, et al. Sea-surface object detection based on electro-optical sensors: A review[J]. IEEE Intelligent Transportation Systems Magazine, 2023, 15(2): 190-216.
- [19] JIANG B, QU R, LI Y, et al. VC-YOLO: Towards real-time object detection in aerial images[J]. Journal of Circuits, Systems and Computers, 2022, 31(8): 2250147.
- [20] 邓天民, 程鑫鑫, 刘金凤, 等. 基于特征复用机制的航拍图像小目标检测算法[J]. 浙江大学学报(工学版), 2024, 58(3): 437-448.
- DENG Tianmin, CHENG Xinxin, LIU Jinfeng, et al. Small target detection algorithm for aerial images based on feature reuse mechanism[J]. Journal of Zhejiang University (Engineering Science), 2024, 58(3): 437-448.
- [21] 邬开俊, 蒲卓. 基于特征信息补充与增强的无人机视角图像目标检测算法[J/OL]. 北京航空航天大学学报, 2024. <https://bhxb.buaa.edu.cn/bhzhk/cn/article/doi/10.13700/j.bh.1001-5965.2024.0190>.
- WU Kaijun, PU Zhuo. UAV-view image object detection algorithm based on feature information supplementation and enhancement[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2024. <https://bhxb.buaa.edu.cn/bhzhk/cn/article/doi/10.13700/j.bh.1001-5965.2024.0190>.
- [22] JIANG L, YUAN B, DU J, et al. MFFSODNet: Multi-scale feature fusion small object detection network for UAV aerial

- images[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 5015214.
- [23] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection[EB/OL]. (2020-06-09). <https://arxiv.org/pdf/1911.09516>
- [24] LI J, WEN Y, HE L. SCConv: Spatial and channel reconstruction convolution for feature redundancy[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, Canada: IEEE, 2023: 6153-6162.
- [25] CHEN J, MAI H S, LUO L, et al. Effective feature fusion network in BIFPN for small object detection[C]//*Proceedings of the 2021 IEEE International Conference on Image Processing (ICIP)*. Anchorage, USA: IEEE, 2021: 699-703.
- [26] VARGA L A, KIEFER B, MESSMER M, et al. SeaDronesSee: A maritime benchmark for detecting humans in open water [C]//*Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision*. Waikoloa, USA: IEEE, 2022: 3686-3696.
- [27] KHANAM R, HUSSAIN M. YOIOv11: An overview of the key architectural enhancements[EB/OL]. (2024-11-07). <https://arxiv.org/abs/2410.17725v1>.
- [28] YU G, CHANG Q, LV W, et al. PP-PicoDet: A better real-time object detector on mobile devices[EB/OL]. (2021-11-01). <https://arxiv.org/abs/2111.00902v1>.
- [29] ZHANG J, LEI J, XIE W, et al. SuperYOLO: Super resolution assisted object detection in multimodal remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 1-15.
- [30] XU J, FAN X, JIAN H, et al. YoloOW: A spatial scale adaptive real-time object detection neural network for open water search and rescue from UAV aerial imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5623115.
- [31] GE Q, DA W, WANG M. MARFPNet: Multi-attention and adaptive reparameterized feature pyramid network for small target detection on water surfaces[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 5039617.
- [32] DU D, ZHU P, WEN L, et al. VisDrone-DET2019: The vision meets drone object detection in image challenge results[C]//*Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision Workshop*. Seoul, South Korea: IEEE, 2019: 213-226.
- [33] JIANG B, QU R, LI Y, et al. VC-YOLO: Towards real-time object detection in aerial images[J]. *Journal of Circuits, Systems and Computers*, 2022, 31(8): 2250147.
- [34] SU J, QIN Y, JIA Z, et al. MPE-YOLO: Enhanced small target detection in aerial imaging[J]. *Scientific Reports*, 2024, 14 (1): 17799.

作者简介:



黄绿娥(1981-),女,副教授,研究方向:深度学习与机器视觉,E-mail:9320080310@jxust.edu.cn。



于晓伟(2001-),男,硕士研究生,研究方向:深度学习与目标检测。



郢化彪(1978-),男,副教授,研究方向:系统建模与嵌入式系统应用。



毛玉婷(1983-),通信作者,女,副研究员,研究方向:人工智能与绿色技术,E-mail:58902885@qq.com。

(编辑:陈璐,王婕)